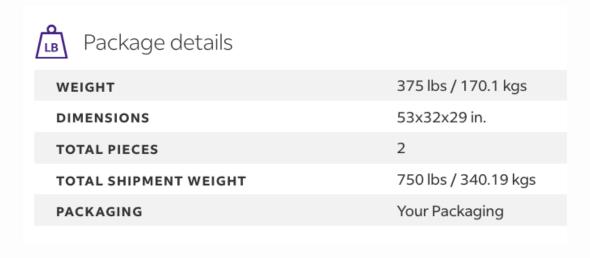
AMD is sending us the two MI300X boxes we asked for. They are in the mail.





It took a bit, but AMD passed my cultural test. I now believe they aren't going to shoot themselves in the foot on software, and if that's true, there's absolutely no reason NVIDIA should be worth 16x more.

CUDA isn't really the moat people think it is, it is just an early ecosystem. tiny corp has a fully sovereign AMD stack, meaning we have rewritten the full stack from the hardware to PyTorch (with the exception of LLVM), and soon we'll port it to the MI300X. You won't even have to use tinygrad proper to use it, tinygrad has a torch frontend now.

Either NVIDIA is super overvalued or AMD is undervalued. If the petaflop gets commoditized (tiny corp's mission), the current situation doesn't make any sense. The hardware is similar, AMD even got the double throughput Tensor Cores on RDNA4 (NVIDIA artificially halves this on their cards, soon market pressure will force them not to).

I'm betting on AMD being undervalued, and that the demand for AI has barely started. With good software, the MI300X should outperform the H100.



Bought in for a quarter million. Long term. It can always dip short term, but check back in 5 years.